

**Space-Time Filtering,  
Sampling and Motion Uncertainty**

**Radu S. Jasinski**

**CMU-RI-TR-88-9**

AD-A199 667

**The Robotics Institute  
Carnegie Mellon University  
Pittsburgh, Pennsylvania 15213**

**June 1988**

**DTIC  
ELECTE  
SEP 14 1988  
S H D**

**© 1988 Carnegie Mellon University**

**This research was supported by the Defense Advanced Research Projects Agency (DoD), monitored by the U.S. Army Engineer Topographic Laboratories under Contract DACA 76-85-C-0002.**

**Address: Center for Automation Research, University of Maryland, College Park, MD 20742**

**DISTRIBUTION STATEMENT A**

**Approved for public release;  
Distribution Unlimited**

**88 9 12 097**

REPORT DOCUMENTATION PAGE				
1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE				
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CMU-RI-TR-88-9			5. MONITORING ORGANIZATION REPORT NUMBER(S) DACA76-85-C-0002	
6a. NAME OF PERFORMING ORGANIZATION The Robotics Institute Carnegie Mellon University		6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U.S. Army Engineer Topographic Laboratories	
6c. ADDRESS (City, State, and ZIP Code) Pittsburgh, PA 15213			7b. ADDRESS (City, State, and ZIP Code)	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION DARPA		8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State, and ZIP Code)			10. SOURCE OF FUNDING NUMBERS	
			PROGRAM ELEMENT NO	PROJECT NO
			TASK NO	WORK UNIT ACCESSION NO
11. TITLE (Include Security Classification) Space-Time Filtering, Sampling and Motion Uncertainty				
12. PERSONAL AUTHOR(S) Radu S. Jasinski				
13a. TYPE OF REPORT Technical		13b. TIME COVERED FROM TO	14. DATE OF REPORT (Year, Month, Day) June 1988	15. PAGE COUNT 36
16. SUPPLEMENTARY NOTATION				
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP		
19. ABSTRACT (Continue on reverse if necessary and identify by block number) We analyse the structure of filters which are space-time oriented. Basically, this paper consists of two parts. In the first, we present the cascade of space-time DOG as an energy filter, discuss its general properties and show how to compute its energy. In the second part, we discuss the consequences of applying the sampling theorem to uniformly translating patterns in the presence of motion uncertainty. It is shown that, for a given motion uncertainty, there exists a bound on the maximum sampling interval, such that for larger values aliasing will occur.				
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL			22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

## Abstract

We analyse the structure of filters which are space-time oriented. Basically, this paper consists of two parts. In the first one, we present the cascade of space-time DOG as an energy filter, discuss its general properties and show how to compute its energy. In the second part, we discuss the consequences of applying the sampling theorem to uniformly translating patterns in the presence of motion uncertainty. It is shown that, for a given motion uncertainty, there exists a bound on the maximum sampling interval, such that for larger values aliasing will occur.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

## 1 Introduction

The problem of the extraction of the optical flow has, in recent years, been treated from a new point of view, that is, through the use of space-time filters [7,4,6,5]. The basic idea behind this method is to extract the optical flow without having to perform any type of operation other than to use a collection of filters which are tuned to different orientations in space-time (or equivalently in the frequency domain). Also, given that the outputs of these filters have been computed, it is necessary to establish a method by which we can determine the value of the estimated optical flow, because the output of a filter tuned to a specific orientation (even if with maximal response) is not enough to extract the optical flow and we have to use a complete set of filters (in the sense that it takes into account all possible orientations). In space-time filtering, we convolve a sequence of images with a (space-time) filter, such that the interval between successive images is small. The minimum temporal interval between successive images is basically dictated by practical considerations, because if it is too small we get little amount of information about the moving pattern from frame to frame. On the other hand, we would like to know what the value of the maximum temporal interval between successive images should be such that we continue to be able to use a filtering approach to the extraction of optical flow

The answer to this question comes by considering the sampling issues involved in this filtering process. As I will show in section 3, if there exists a certain degree of motion uncertainty, then the maximum sampling interval, is fixed by this motion uncertainty. This means that there exists a (non-linear) relationship between the motion uncertainty and the maximum sampling interval.

The procedure of using a collection of filters to extract optical flow corresponds, in a general sense, to a signal processing approach, which is mainly concerned with the extraction of information about the original signal, in the presence of noise. It involves the construction of filters, if possible optimal ones, parameter estimation and the analysis of sampling issues.

On the other hand, in the feature based approach to the extraction of optical flow [3] it is necessary to, previously to the actual computation of the optical flow, extract edges (zero crossings) which

have to be matched in successive frames. If the temporal interval between successive images is large and the number of edges to be matched between frames is not high, then a contour-based approach can be successful in extracting the optical flow. But, on the other hand, if the number of edges is large, the amount of mismatches can lead to a high rate of error, and as a consequence of this to a wrong estimation of the optical flow.

It is therefore important to be able to detect moving features and extract their optical flow in the presence of noisy data and imprecise measurements. Depending on the spatial complexity of information available at each image, in the temporal sequence of images, it can be more reliable to use a filtering approach, especially for the case in which the interval between these images is small and there exists a high spatial content of information (which makes a feature matching approach highly unstable).

In this paper we discuss, in section 2, the issue of extracting the optical flow through feature or intensity based approaches versus space-time filtering, and present the space-time DOG cascade as an energy filter. In section 3 we analyse sampling issues which apply for uniformly translating patterns in the presence of noise (motion uncertainty). Finally, we draw conclusions in section 4, and make an analogy between the long and short-range processes of motion extraction in the human visual system and the feature-based and space-time filtering methods in Computer Vision.

## 2 Space-time filtering

### 2.1 Extraction of the optical flow in intensity and feature-based approach

The extraction of the optical flow field from the intensity variations in the image plane has been treated until very recently, in Computer Vision, as a feature or intensity-based problem. In the feature-based approach we have to detect relevant features, such as edges, from a pair of successive images (in a temporal sequence of images), and afterwards perform a matching of corresponding elements, so

that, as a result of this procedure, we assign a specific value of the optical flow to the corresponding elements.

This method has to overcome two major problems:

1. The correspondence problem
2. The aperture problem.

The correspondence problem [1] addresses the question of how to assign the same identity for elements which appear in temporal succession of images. The correspondence, or matching, can be computed in different ways, depending, in part, on the temporal interval between successive images. If this interval is small, the correspondence between features can be performed through a set of local operations over elements which are spatially close to each other. One of these operations [1,17] consists in the minimization of the distance a set of elements takes to travel from one image to its successive one. On the other hand, if this temporal interval is large, it is more likely that a more global type of operation for the matching of features has to be implemented. In general, the matching of corresponding elements in successive images can be unstable, due to noise in the image, and also computationally expensive if the number of features to be matched is large.

In respect to the aperture problem, which states that it is not possible to measure both components of the optical flow field given a small aperture in the image, we have to introduce additional constraints into the model describing the extraction of optical flow, so as to make it possible to obtain the full optical flow field. Actually, given a small aperture, we are only able to measure the normal component (to the gradient of the intensity) of optical flow field, while its tangential component remains undetermined. As one example of the solution to the aperture problem, we can mention the area-based [2] formulation which assumes the use of a smoothness term, in addition to the intensity continuity equation, represented by the sum of the squares of the spatial derivatives of the optical flow field components. Another example is given by the contour-based [3] approach, where the constraint is represented by the gradient in respect to the arc length along the intensity gradient of the optical

flow field, in addition to the difference between the normal component of the optical flow and its measured value.

Both, the correspondence and the aperture problem, involve in practice a certain amount of arbitrariness in terms of having to choose a set of constraints which enable us to extract the full optical flow. It is therefore desirable to be able to eliminate the necessity of having to cope with both of these problems. The method of space-time filtering does this, in part, by eliminating altogether the necessity of the use of the correspondence problem. In respect to the aperture problem the solution given by Heeger [5] consists in modeling the image flow as (locally) purely translational, so that the optical flow is extracted by fitting a plane to the energy of the filter. This is equivalent to the computation of both components of the optical flow field, because for translational motion the support in the frequency domain is given by a plane whose orientation is a function of the velocity vector.

## 2.2 Space-time oriented filters

Space-time filtering consists, basically, in the convolution of a temporal sequence of images (closely displaced) with a (space-time) filter. The most important aspect of space-time filtering lies in the fact that, if we consider an uniformly translating pattern, we are able to select a specific velocity by using (space-time) oriented filters [6].

Let us take the example of one-dimensional motion (in the  $x$  direction). If we analyse the picture which is generated in space-time by an uniformly translating pattern (through a cross-section parallel the  $x$ - $t$  plane), then we can conclude that the orientation of the individual elements (like lines or stripes) is intrinsically determined by the velocity of the pattern (the slope of a line in the EPI plane is equal to the velocity of the feature associated to it). A very interesting example of this kind of relationship between (space-time) orientation and velocity is described by the epipolar plane images (EPIs) created by Bolles and Baker [8] for the case of a camera moving (perpendicularly to the direction of motion) in a static environment. There, at a given EPI, we are able to track the temporal evolution of each image element (at a fixed height), and this is described by a straight line.

Once we know that, for uniform translation, the space-time evolution of image elements is given by straight lines, in order to select a specific velocity (optical flow), we can use (space-time) oriented filters [6,7]. A particularly important aspect of this analysis comes from the fact that, for an uniformly translating pattern, the support of the contrast function in the frequency domain is given by a plane (or a line for the case of one-dimensional motion) [7,10] passing through the origin of the coordinate system. In terms of space-time filtering, this means that, in order to select a specific velocity of an uniformly translating pattern, we have to tune the filter to the orientation in the frequency domain which gives the highest response.

The use of directionally selective (space-time) filters pose a limitation in the sense that they are phase sensitive [6]. This means that, depending on the alignment between the space-time configuration of moving patterns and the filter shape, we can get different results: the filtered output may oscillate or vary between positive and negative values. A solution to this problem is given by computing the energy (power spectrum) of the filter output. The energy of a convolved signal is independent of any phase problem, and for the case of an uniformly translating pattern its output is constant.

If, for example, we compute the energy associated to a space-time Gabor filter [5] convolved with an arbitrary function, then the final result will not oscillate or depend on any phase factor. This leads to the concept of space-time oriented filters as energy filters, which, with the assumption of random textured images and Parseval's theorem made it possible for Heeger [5] to, analytically, predict the energy associated to a particular space-time oriented pattern.

### 2.3 Space-time Difference-of-Gaussian (DOG) cascade as an energy filter

Space-time filtering, either through energy filters or cascades, is primarily concerned with the processing of a temporal sequence of images, such that the interval between successive images is small. On one hand, the work of Heeger [5] showed us that it is possible to obtain a dense image flow



map by using a collection of twelve space-time Gabor filters, each tuned to a different direction in space-time. The space-time Gabor filter is parametrized by three (gaussian) filter sizes ( $\sigma_x$ ,  $\sigma_y$  and  $\sigma_t$ ), in addition to the (three) sine or cosine space-time frequencies (whose relative ratios correspond to different orientations in space-time). It would be desirable to have a broader space-time tuning capability, as, for example, in the case of the cascaded filters proposed by Fleet and Jepson [9]. They proposed the construction of space-time oriented filters in terms of cascades of the CS filter. The CS filter is defined as the difference of spatial gaussians which are each multiplied by a temporal exponentially decaying function, corresponding to a temporal center (C)-surround (S) model, in analogy to biological systems, plus, a temporal delay term embodied in the S part. The space-time orientation is obtained by convolving the CS filter with a sum of (space-time) Dirac distributions, each centered at a specific location in space-time so that the result is a oriented pattern. The use of layered cascades of the CS filter improves the orientation specificity of the filters, as shown by Fleet and Jepson [9]. In respect to its tuning capabilities, these layered cascades of the CS filter, are able, in addition to their specific orientation, to select features moving at high or low speed by adjusting the ratio of the spatial or temporal filter sizes to one, respectively.

We would like to use a filter which exhibits a wide range of space-time tuning and can also be used to extract the image flow as an energy filter. The simplest fusion of these two aspects is exhibited by the space-time DOG filter, used in cascade. In fact, if we substitute the temporal exponential decay term in the CS filter by a temporal gaussian, and eliminate the temporal delay, we get a space-time DOG. The number of parameters of this filter is equal 9, where 4 correspond to the center and surround filter sizes (the spatial filter sizes are assumed to be equal), spatial and temporal offsets make up 3 parameters, plus the center and surround multiplicative constants. The only reason for not using the CS cascade filter of Fleet and Jepson directly as an energy filter comes from the fact that the energy expression turns out to be more complex than that of the space-time DOG cascade because it has a linear temporal exponential decay, whereas for the DOG filter the temporal decay is gaussian, thus making it easier to perform the temporal integral in order to get the energy expression.

We should remind ourselves that the space-time DOG and Gabor filters are non-causal, as a consequence of the Paley-Wiener theorem [12] which states that, if a temporal filter  $f(t)$  has a square

integrable Fourier transform  $\tilde{f}(w)$  and satisfies the relation

$$\int_{-\infty}^{\infty} \frac{|\ln |\tilde{f}(w)||}{1+w^2} dw < \infty, \quad (2.1)$$

then  $f(t)$  is causal. The CS filter, on the contrary, due to its linear exponential temporal decay, is a causal filter.

The space-time DOG filter is given by the following expression

$$\begin{aligned} D(x, y, t) = & A_c \exp(-(x^2 + y^2)/(2\sigma_c^2) - t^2/(2\mu_c^2)) \\ & - A_s \exp(-(x^2 + y^2)/(2\sigma_s^2) - t^2/(2\mu_s^2)), \end{aligned} \quad (2.2)$$

where  $\sigma_c$  ( $\sigma_s$ ) and  $\mu_c$  ( $\mu_s$ ) are the center (surround) spatial and temporal filter sizes respectively, while  $A_c$  and  $A_s$  are adjustable parameters (used in the discrete version of the filter to tune the sum of all elements of the mask to zero). Its Fourier transform is given by

$$\begin{aligned} D(\vec{k}, w) = & A_c \exp(-(\vec{k}^2 \sigma_c^2/2 + w^2 \mu_c^2/2)) \\ & - A_s \exp(-(\vec{k}^2 \sigma_s^2/2 + w^2 \mu_s^2/2)), \end{aligned} \quad (2.3)$$

where the spatial and temporal frequencies are respectively given by  $\vec{k}$  ( $\vec{k} = (k_x, k_y)$ ) and  $w$ .

A cascade of filters corresponds to applying, in succession, a set of linear filters, to a collection of signals [9], such that the interval between their successive positions of highest magnitude is measured by the offset. In the case of space-time filtering these offsets have a spatial as well as a temporal part. Also, they can occur in a set of layers, where each layer corresponds to a different collection of space-time offsets.

Let us define, analogously to Fleet and Jepson [9], the one-layer cascade by the expression

$$C(x, y, t) = D(x, y, t) * E(x, y, t), \quad (2.4)$$

where

$$E(x, y, t) = \frac{1}{2}\delta(x, y, t) + \frac{1}{4}\delta(x + \xi_x, y + \xi_y, t + \tau) + \frac{1}{4}\delta(x - \xi_x, y - \xi_y, t - \tau), \quad (2.5)$$

$\delta(\cdot)$  is the Dirac delta (distribution),  $D(x, y, t)$  the space-time DOG as given by formula (2.2) and  $*$  the convolution operation. In the frequency domain this cascade is given by

$$\tilde{C}(\vec{k}, w) = \tilde{D}(\vec{k}, w) \tilde{E}(\vec{k}, w) \quad (2.6)$$

where

$$\begin{aligned} \tilde{E}(\vec{k}, w) &= \frac{1}{2} + \frac{1}{4} [\exp(i(\vec{k} \cdot \vec{\xi} + w\tau)) + \exp(-i(\vec{k} \cdot \vec{\xi} + w\tau))] \\ &= \frac{1}{2} [1 + \cos(\vec{k} \cdot \vec{\xi} + w\tau)], \end{aligned} \quad (2.7)$$

and  $\tilde{D}(\vec{k}, w)$  is given by (2.3).

By increasing (decreasing) the offset values of, for example, the one-layer cascade we get more (less) specificity to velocity. This can be observed by comparing Figures 1 and 2, or their respective Fourier transform, Figures 3 and 4. We fix  $U_c = 1.0$ ,  $U_s = 3.0$ ,  $M_c = 1.0$ ,  $M_s = 3.0$ ,  $A_c = 1.0$  and  $A_s = 1.0$ .

For Figure 1 we have  $\xi_x = 0.77$ ,  $\xi_y = 0$  and  $\tau = 2.89$ , whereas for Figure 2,  $\xi_x = 0.52$ ,  $\xi_y = 0$  and  $\tau = 1.93$ , which correspond to a slope of 15 deg in the  $x$ - $t$  plane (or 0.26 pixels per frame). If we inspect Figures 3 and 4 it becomes clear that for larger offsets (Figure 3) we get more tuning to velocity, although more ringing [9] (due to aliasing of adjacent patterns), for small velocities. A way by which we get less ringing and more velocity specificity, as described by Fleet and Jepson [9], is to build cascades out of more than one layer. For example, a two layered cascade is constructed by convolving two one-layer cascades, each with a different collection of offset values, that is

$$C(x, y, t) = C_1(x, y, t) * C_2(x, y, t), \quad (2.8)$$

where

$$C_1(x, y, t) = D(x, y, t) * E_1(x, y, t), \quad (2.9)$$

$$C_2(x, y, t) = D(x, y, t) * E_2(x, y, t), \quad (2.10)$$

$$E_1(x, y, t) = \frac{1}{2}\delta(x, y, t) + \frac{1}{4}\delta(x + \xi_x^1, y + \xi_y^1, t + \tau^1) + \frac{1}{4}\delta(x - \xi_x^1, y - \xi_y^1, t - \tau^1), \quad (2.11)$$

and

$$E_2(x, y, t) = \frac{1}{2}\delta(x, y, t) + \frac{1}{4}\delta(x + \xi_x^2, y + \xi_y^2, t + \tau^2) + \frac{1}{4}\delta(x - \xi_x^2, y - \xi_y^2, t - \tau^2). \quad (2.12)$$

For this two-layer cascade, Figure 5, and its Fourier transform Figure 6, we can observe (we use the same space-time scale as for the one-layer cascade Figures) that there occurs much less ringing (Figure 6) and its space-time shape exhibits a broader (also narrower), support at the particular orientation for which it is tuned.

In general, irrespective of the set of parameters that we choose for the filter, there always exists a specific amount of directional uncertainty which is a consequence of the fact that the filter response is not perfectly tuned to a particular orientation. This is a consequence of the fact that, in addition to the response of the filter to the particular orientation for which is tuned, there exists a non-zero response to a restricted range of orientations in its neighborhood. For example, in the case of one-dimensional motion (parallel to the  $x$  axis) of a given image pattern, in order to filter the specific direction (in the  $x$ - $t$  plane or, equivalently, in the frequency domain) associated to its velocity, we should use a filter which exhibits its support at a given orientation and is zero otherwise. In practice, we will only be able to select a given orientation inside a cone, such that its aperture is proportional to the motion uncertainty. This is a consequence of the fact that any (real) filter will not only select the particular direction for which it was designed, but also adjacent directions inside a fixed aperture. As a result of this, there will always result a motion uncertainty, and consequently, this will affect (space-time) the sampling properties of the filter. This issue will be discussed in detail in the next section.

The energy (power spectrum) associated to the one-layer cascade is given by

$$\begin{aligned} & \int_{-\infty}^{\infty} d\vec{k} \int_{-\infty}^{\infty} d\omega |\tilde{C}(\vec{k}, \omega)|^2 \\ &= \int_{-\infty}^{\infty} d\vec{k} \int_{-\infty}^{\infty} d\omega \{ \tilde{D}^2(\vec{k}, \omega) [1 + 2 \cos(\vec{k} \cdot \vec{\xi} + \omega \tau) + \cos^2(\vec{k} \cdot \vec{\xi} + \omega \tau)] \}. \end{aligned} \quad (2.13)$$

Since we assume only translational motion, in which case it holds that

$$\omega = \vec{k} \cdot \vec{v} \quad (2.14)$$

where  $\vec{v}$  is the velocity field, we can rewrite the previous energy expression in the following form

$$\begin{aligned} & \int_{-\infty}^{\infty} d\vec{k} |\tilde{C}(\vec{k}, \vec{k} \cdot \vec{v})|^2 \\ &= \int_{-\infty}^{\infty} d\vec{k} \{ \tilde{D}^2(\vec{k}, \vec{k} \cdot \vec{v}) [ \frac{3}{2} + [ \exp(i\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) + \exp(-i\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) ] \} \end{aligned}$$

$$+ \frac{1}{4} [\exp(2i\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) + \exp(-2i\vec{k} \cdot (\vec{\xi} + \vec{v}\tau))] \}. \quad (2.15)$$

As a next step we want to develop the expression of the energy (2.15), by performing the integral over  $\vec{k}$ . In this respect, it is useful to notice that formula (2.15) contains the following algebraic expression

$$I(\vec{\xi}, \tau, a, \vec{v}) = \int_{-\infty}^{\infty} d\vec{k} \{ D^2(\vec{k}, \vec{k} \cdot \vec{v}) [\exp(ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) + \exp(-ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau))] \}, \quad (2.16)$$

where  $a$  can be any positive integer number. If we use the definition of  $\tilde{D}(\vec{k}, w)$  and the constraint (2.14), then

$$\begin{aligned} D^2(\vec{k}, \vec{k} \cdot \vec{v}) &= A_c^2 \exp(-k_x^2(\sigma_c^2 + v_x^2\mu_c^2) - k_y^2(\sigma_c^2 + v_y^2\mu_c^2) - 2k_xk_yv_xv_y\mu_c^2) \\ &\quad + A_s^2 \exp(-k_x^2(\sigma_s^2 + v_x^2\mu_s^2) - k_y^2(\sigma_s^2 + v_y^2\mu_s^2) - 2k_xk_yv_xv_y\mu_s^2) \\ &\quad - 2A_cA_s \exp(-k_x^2(\sigma_c^2 + \sigma_s^2 + v_x^2(\mu_c^2 + \mu_s^2)) - k_y^2(\sigma_c^2 + \sigma_s^2 + v_y^2(\mu_c^2 + \mu_s^2))) \\ &\quad \times \exp(-2k_xk_yv_xv_y(\mu_c^2 + \mu_s^2)). \end{aligned} \quad (2.17)$$

Now, by inserting (2.17) into expression (2.16), we get

$$I(\vec{\xi}, \tau, a, \vec{v}) = F_1(a, \vec{\xi}, \vec{v}, \tau) + F_2(a, \vec{\xi}, \vec{v}, \tau) + F_3(a, \vec{\xi}, \vec{v}, \tau), \quad (2.18)$$

where

$$\begin{aligned} F_1(a, \vec{\xi}, \vec{v}, \tau) &= A_c^2 \int_{-\infty}^{\infty} d\vec{k} \{ [\exp(ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) + \exp(-ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau))] \\ &\quad \times \exp(-k_x^2(\sigma_c^2 + v_x^2\mu_c^2) - k_y^2(\sigma_c^2 + v_y^2\mu_c^2) - 2k_xk_yv_xv_y\mu_c^2) \}. \end{aligned} \quad (2.19)$$

$$\begin{aligned} F_2(a, \vec{\xi}, \vec{v}, \tau) &= A_s^2 \int_{-\infty}^{\infty} d\vec{k} \{ [\exp(ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) + \exp(-ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau))] \\ &\quad \times \exp(-k_x^2(\sigma_s^2 + v_x^2\mu_s^2) - k_y^2(\sigma_s^2 + v_y^2\mu_s^2) - 2k_xk_yv_xv_y\mu_s^2) \}. \end{aligned} \quad (2.20)$$

and

$$\begin{aligned} F_3(a, \vec{\xi}, \vec{v}, \tau) &= -2A_cA_s \int_{-\infty}^{\infty} d\vec{k} \{ [\exp(ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau)) + \exp(-ia\vec{k} \cdot (\vec{\xi} + \vec{v}\tau))] \\ &\quad \times \exp(-k_x^2(\sigma_c^2 + \sigma_s^2 + v_x^2(\mu_c^2 + \mu_s^2)) - k_y^2(\sigma_c^2 + \sigma_s^2 + v_y^2(\mu_c^2 + \mu_s^2))) \\ &\quad \times \exp(-2k_xk_yv_xv_y(\mu_c^2 + \mu_s^2)) \}. \end{aligned} \quad (2.21)$$

Finally, if we use the (gaussian) integral formula

$$\begin{aligned} & \int_{-\infty}^{\infty} d\vec{k} \exp(i(\varphi_x \varphi_y)) \begin{pmatrix} k_x \\ k_y \end{pmatrix} + -\frac{1}{2}(k_x k_y) \Lambda^{-1} \begin{pmatrix} k_x \\ k_y \end{pmatrix} \\ &= [4\pi^2 \det \Lambda]^{\frac{1}{2}} \exp(-\frac{1}{2}(\varphi_x \varphi_y) \Lambda \begin{pmatrix} \varphi_x \\ \varphi_y \end{pmatrix}), \end{aligned} \quad (2.22)$$

for an arbitrary  $(2 \times 2)$  matrix  $\Lambda$ , then  $F_1$  will be given by

$$\begin{aligned} F_1(a, \bar{\xi}, \bar{v}, \tau) &= \frac{2\pi}{\sqrt{2}} [(\sigma_c^2 + \mu_c^2 v_x^2)(\sigma_c^2 + \mu_c^2 v_y^2) - (v_x v_y \mu_c^2)^2]^{-\frac{1}{2}} \\ &\quad \times \exp(-\frac{1}{4} a^2 [(\sigma_c^2 + \mu_c^2 v_x^2)(\sigma_c^2 + \mu_c^2 v_y^2) - (v_x v_y \mu_c^2)^2]^{-1}) \\ &\quad \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \begin{pmatrix} \sigma_c^2 + \mu_c^2 v_x^2 & -v_x v_y \mu_c^2 \\ -v_x v_y \mu_c^2 & \sigma_c^2 + \mu_c^2 v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \end{aligned} \quad (2.23)$$

Analogously,  $F_2$  and  $F_3$  are given by a similar expression, if we substitute  $\sigma_c^2, \mu_c^2$  by  $\sigma_s^2, \mu_s^2$  and  $\tau_c^2 + \sigma_s^2, \mu_c^2 + \mu_s^2$ , respectively.

The complete expression for the energy is given by

$$\begin{aligned} & \int_{-\infty}^{\infty} d\vec{k} |\tilde{C}(\vec{k}, \vec{k} \cdot \vec{v})|^2 \\ &= \frac{2\pi}{\sqrt{2}} \{ A_c^2 [(\sigma_c^2 + \mu_c^2 v_x^2)(\sigma_c^2 + \mu_c^2 v_y^2) - (v_x v_y \mu_c^2)^2]^{-\frac{1}{2}} [\frac{3}{2} \\ &\quad + \exp(-\frac{1}{4} [(\sigma_c^2 + \mu_c^2 v_x^2)(\sigma_c^2 + \mu_c^2 v_y^2) - (v_x v_y \mu_c^2)^2]^{-1}) \\ &\quad \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \begin{pmatrix} \sigma_c^2 + \mu_c^2 v_x^2 & -v_x v_y \mu_c^2 \\ -v_x v_y \mu_c^2 & \sigma_c^2 + \mu_c^2 v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \\ &\quad + \frac{1}{4} \exp(-[(\sigma_c^2 + \mu_c^2 v_x^2)(\sigma_c^2 + \mu_c^2 v_y^2) - (v_x v_y \mu_c^2)^2]^{-1}) \\ &\quad \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \begin{pmatrix} \sigma_c^2 + \mu_c^2 v_x^2 & -v_x v_y \mu_c^2 \\ -v_x v_y \mu_c^2 & \sigma_c^2 + \mu_c^2 v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \} \\ &\quad + A_s^2 [(\sigma_s^2 + \mu_s^2 v_x^2)(\sigma_s^2 + \mu_s^2 v_y^2) - (v_x v_y \mu_s^2)^2]^{-\frac{1}{2}} [\frac{3}{2} \\ &\quad + \exp(-\frac{1}{4} [(\sigma_s^2 + \mu_s^2 v_x^2)(\sigma_s^2 + \mu_s^2 v_y^2) - (v_x v_y \mu_s^2)^2]^{-1}) \end{aligned}$$

$$\begin{aligned}
& \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \begin{pmatrix} \sigma_s^2 + \mu_s^2 v_x^2 & -v_x v_y \mu_s^2 \\ -v_x v_y \mu_s^2 & \sigma_s^2 + \mu_s^2 v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \\
& + \frac{1}{4} \exp(-[(\sigma_s^2 + \mu_s^2 v_x^2)(\sigma_s^2 + \mu_s^2 v_y^2) - (v_x v_y \mu_s^2)^2]^{-1}) \\
& \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \begin{pmatrix} \sigma_s^2 + \mu_s^2 v_x^2 & -v_x v_y \mu_s^2 \\ -v_x v_y \mu_s^2 & \sigma_s^2 + \mu_s^2 v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \\
& - 2A_c A_s [(\sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_x^2)(\sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_y^2) - (v_x v_y (\mu_c^2 + \mu_s^2))^2]^{-\frac{1}{2}} \left[ \frac{3}{2} \right. \\
& + \exp(-\frac{1}{4}[(\sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_x^2)(\sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_y^2) - (v_x v_y (\mu_c^2 + \mu_s^2))^2]^{-1}) \\
& \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \\
& \times \begin{pmatrix} \sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_x^2 & -v_x v_y (\mu_c^2 + \mu_s^2) \\ -v_x v_y (\mu_c^2 + \mu_s^2) & \sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \\
& + \frac{1}{4} \exp(-[(\sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_x^2)(\sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_y^2) - (v_x v_y (\mu_c^2 + \mu_s^2))^2]^{-1}) \\
& \times ((\xi_x + v_x \tau)(\xi_y + v_y \tau)) \\
& \times \begin{pmatrix} \sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_x^2 & -v_x v_y (\mu_c^2 + \mu_s^2) \\ -v_x v_y (\mu_c^2 + \mu_s^2) & \sigma_c^2 + \sigma_s^2 + (\mu_c^2 + \mu_s^2)v_y^2 \end{pmatrix} \begin{pmatrix} \xi_x + v_x \tau \\ \xi_y + v_y \tau \end{pmatrix} \Big] \Big\}. \quad (2.24)
\end{aligned}$$

If we want to extract the optical flow field by using a set of energy filters, each tuned to a different orientation in space-time, then we are confronted with another source of motion uncertainty. This comes from the fact that we have to determine the optical flow field, given the output of a number of energy filters (with different orientations). For example, in Heeger's approach the estimated field  $(v_x, v_y)$  minimizes a cost function, which consists in the sum of the difference between the measured (motion) energy and its predicted value, over all twelve filters. This means that there will always exist a non-zero contribution from filters which do not correspond to the right orientation, due to an overlap in the shape of neighboring filters. If we wish to reduce the uncertainty in the motion estimate because of neighboring interaction among filters, we have to enhance the orientation specificity of each filter (thus leading to less lateral overlap). But this has the consequence that, for a fixed number of filters, some orientation (mainly corresponding to the orientations between that of neighboring filters) will not be able to be selected any more. So we are faced with a trade-off between being able to select a specific orientation in space-time, with a minimum of uncertainty, and the minimum

numbers of filters necessary to span all orientations.

The method of optical flow extraction used by Heeger [5], although it is able to determine the optical flow for a collection of different types of moving patterns, contains some limitations which should be mentioned, that is:

1. It assumes that all images can be modeled as (locally) random patterns
2. In order to be able to use Parseval's theorem, it is necessary to approximate the expression of the energy
3. The optimization procedure, which has to be performed at each image pixel, is computationally very expensive.

In particular, the issue of approximating the integral in Parseval's theorem, leads to errors in the estimated value of the optical flow in regions where it is discontinuous, thus making it difficult to use the estimated value as input for the operation of region segmentation. This and other questions will be discussed in another paper [11].

### 3 Space-time sampling

In the previous section I discussed the question of extracting the optical flow by using space-time filters, considered as energy filters. Also, I proposed the use of cascades of space-time filters like the ones constructed by Fleet and Jepson as energy filters, which can be accomplished by substituting the temporal exponential by a gaussian. A consequence of adopting a filtering approach to the extraction of the optical flow is the fact that it is necessary to sample the filter, or more specifically, to perform a space-time sampling of the filter. The temporal sampling issue is very clearly determined by the fact that the temporal interval between successive images used in space-time filtering, although small,



is finite. A question which is naturally raised in this context is in respect to how much can we (temporally) undersample the filter, or in a more complete statement, the convolution of the image sequence with the space-time filter, so that we are still able to reconstruct the original signal. For uniformly translating patterns, the spatial and temporal sampling ratios are not independent, and, as it is shown next, if there exists a certain amount of motion uncertainty, then there exists a maximum sampling interval, in either space or time, such that aliasing does not occur. This maximum sampling interval is shown to be a (non-linear) function of the motion uncertainty.

Initially, I will describe very succinctly, for one-dimensional functions, the sampling theorem and generalize it to three-dimensions (two spatial and one temporal). Next, I show that for an uniformly translating pattern it is only necessary to sample in either the spatial or temporal variables. Finally, I relate motion uncertainty with the maximum sampling interval such that there is no aliasing.

The sampling theorem [12] gives us a mathematical formulation for the reconstruction of a continuous function in terms of a collection of samples of this function, over a specific domain. If we deal with real signals, on the other hand, there is always a certain amount of under or oversampling depending on the specific architecture of the filters being used. In particular, for the case of undersampling (where the spatial or temporal sampling rate is larger than the one established by the sampling theorem - the Nyquist rate), we have to deal with the aliasing problem. The degree of aliasing which is permitted (so that it still is possible to reconstruct the original function, modulo small distortions) depends not only on the filter characteristics but also on the type of data being filtered.

Let us start with one-dimensional signals, represented by the function  $f(x)$ . We obtain a sample of  $f(x)$ ,  $f_s(x)$ , by multiplying it by a (infinite) sum of (Dirac) delta distributions, such that the sample points are equidistant (by  $\rho_x$ ). The sample function  $f_s(x)$  is given by

$$f_s(x) = f(x) \Sigma(x, \rho_x), \quad (3.1)$$

where

$$\Sigma(x, \rho_x) = \sum_{n_x = -\infty}^{\infty} \delta(x - n_x \rho_x). \quad (3.2)$$

In the frequency domain, (3.1) is represented by the convolution

$$\tilde{f}_s(k_x) = \tilde{f}(k_x) * \tilde{\Sigma}(k_x, \rho_x), \quad (3.3)$$

with

$$\tilde{\Sigma}(k_x, \rho_x) = \frac{1}{\rho_x} \sum_{n_x=-\infty}^{\infty} \delta(k_x - \frac{n_x}{\rho_x}). \quad (3.4)$$

If we assume that  $\tilde{f}(k_x)$  is band-limited ( $\tilde{f}(k_x)$  is zero for  $|k_x| > L_x$ ), then it is easy to check that, unless  $\rho_x \leq \frac{1}{2L_x}$ , there will exist a region where the adjacent lobes overlap, which is a signal of undersampling, and as a consequence of this we have the aliasing phenomenon. In order to avoid this from happening, we multiply formula (3.3) by a function  $H(k_x)$ , as for example the ideal low-pass filter (which is 1 for  $|k_x| \leq L_x$ , and 0 otherwise), as a result of which (3.3) reduces to  $\tilde{f}(k_x)$ . This has the consequence that  $f(x)$  can be exactly recovered from its samples. We can synthesize this result, by stating that, if  $\tilde{f}(k_x)$  is band-limited and has no singularities at its extremities ( $k_x = \pm L_x$ ), then

$$f(x) = \sum_{n_x=-\infty}^{\infty} f(\frac{n_x}{2L_x}) \text{sinc}(2L_x(x - \frac{n_x}{2L_x})), \quad (3.5)$$

where

$$\text{sinc}(x) = \frac{\sin \pi x}{\pi x}, \quad (3.6)$$

which is a version of the sampling theorem [13].

We can generalize the sampling theorem to three-dimensional functions. So, given that  $f(x, y, t)$  is a (space-time) function and  $\tilde{f}(k_x, k_y, w)$  its Fourier transform ( $k_x, k_y$  and  $w$  are the Fourier variables associated to  $x, y$  and  $t$ ),  $\tilde{f}$  is zero for  $|k_x| > L_x$ ,  $|k_y| > L_y$  and  $|w| > L_t$  and it does not have singularities at  $|k_x| = L_x$ ,  $|k_y| = L_y$  and  $|w| = L_t$ , then, by the sampling theorem

$$\begin{aligned} f(x, y, t) = & \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} \sum_{n_t=-\infty}^{\infty} f(\frac{n_x}{2L_x}, \frac{n_y}{2L_y}, \frac{n_t}{2L_t}) \text{sinc}(2L_x(x - \frac{n_x}{2L_x})) \\ & \times \text{sinc}(2L_y(y - \frac{n_y}{2L_y})) \text{sinc}(2L_t(t - \frac{n_t}{2L_t})). \end{aligned} \quad (3.7)$$

The case of translational motion [14,15], in which case it holds that

$$\tilde{f}(k_x, k_y, w) = \tilde{f}(k_x, k_y) \delta(w - \vec{k} \cdot \vec{v}), \quad (3.8)$$

(which is equivalent to say that  $w$  is different from zero only on the plane determined by  $\vec{k} \cdot \vec{v}$ ), the sampling theorem reads

$$f(x - v_x t, y - v_y t) = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} f\left(\frac{n_x}{2L_x}, \frac{n_y}{2L_y}\right) \text{sinc}\left(2L_x\left(x - v_x t - \frac{n_x}{2L_x}\right)\right) \times \text{sinc}\left(2L_y\left(y - v_y t - \frac{n_y}{2L_y}\right)\right). \quad (3.9)$$

This means that we only need to sample  $f(x, y)$ , at the Nyquist rate, in terms of its spatial variables. Another way to understand this issue is given in terms of a Fourier analysis, which, as a matter of simplicity, we apply for the two-dimensional case ( $x$ - $t$  space). We know that for pure translation, because of formula (3.8), the sampled function  $\tilde{f}_s(k_x, w)$  (analogously to (3.3)) is given by

$$\tilde{f}_s(k_x, w) = \tilde{f}(k_x) \delta(w - k_x v_x) * \left[ \sum_{n_x=-\infty}^{\infty} 2L_x \delta(k_x - 2n_x L_x) \sum_{n_t=-\infty}^{\infty} 2L_t \delta(w - 2n_t L_t) \right], \quad (3.10)$$

which can be rewritten in the form

$$\tilde{f}_s(k_x, w) = \sum_{n_x=-\infty}^{\infty} \sum_{n_t=-\infty}^{\infty} \int dk'_x \int dw' \tilde{f}(k'_x) \delta(w' - k'_x v_x) \times 2L_x \delta(k_x - k'_x - 2n_x L_x) 2L_t \delta(w - w' - 2n_t L_t). \quad (3.11)$$

By using that

$$\int dx \delta(x - a) \delta(x - b) = \delta(a - b), \quad (3.12)$$

in the integral over  $w'$ , and

$$\int dx f(x) \delta(x - a) \delta(x - b) = f(a) \delta(a - b), \quad (3.13)$$

in the  $k'_x$  integral, we have that (3.11) results in

$$\tilde{f}_s(k_x, w) = \sum_{n_x=-\infty}^{\infty} [\tilde{f}(k_x - 2n_x L_x) 4L_x L_t \sum_{n_t=-\infty}^{\infty} \delta(w - k_x v_x - 2(n_x v_x L_x - n_t L_t))]. \quad (3.14)$$

We can conclude that, if we start by assuming that  $f(x - v_x t)$  is sampled independently in its spatial and temporal variables, then, due to the constraint of uniform translation, we are led to conclude that

we only need to sample in the spatial (temporal) variable. So, we can simplify equation (3.10) to the following form

$$\tilde{f}_s(k_x, w) = \tilde{f}(k_x) \delta(w - k_x v_x) * \left[ \sum_{n_x=-\infty}^{\infty} 2L_x \delta(k_x - 2n_x L_x) \right], \quad (3.15)$$

or, by expanding the convolution we get

$$\tilde{f}_s(k_x, w) = \sum_{n_x=-\infty}^{\infty} \tilde{f}(k_x - 2n_x L_x) 2L_x \delta(w - v_x(k_x - 2n_x L_x)), \quad (3.16)$$

which leads us to the two-dimensional version of equation (3.9). The expression (3.16) is identical to the one describing the Burr's experiment [15] which consists in sampling in space, at a fixed temporal interval, a pattern which moves at constant rate.

For illustration, if we consider a (space-time) band-limited function which describes an uniformly translational motion, then, by the constraint (3.8) its support (in frequency domain) is given by a line segment, as it is shown in Figure 7. Its sampled version, satisfying equation (3.16) with  $M_x = 2L_x$ , consists of a collection of replicas of the original line segment, which are uniformly sampled at intervals of  $M_x$  (see Figure 8).

From this we can deduce that, once the support of  $\tilde{f}(k_x)$  is defined by the straight lines whose slope is given by  $v_x$ , its sampling rate is equal to the spatial sampling rate (or equivalently to the temporal sampling). The function  $f(x, t)$  (which is identical to  $f(x + v_x t)$ ) can be reconstructed from  $\tilde{f}_s(k_x, w)$  by applying a filter which has a support parallel to the line  $w = k_x v_x$ , and more than this, as it is shown by Crick *et al.* [14], this support can be reduced to an infinitesimally narrow strip, as long as there is no motion uncertainty. This means that, for the case of translational motion, we can increase the sampling rate  $\rho_x$  as much as we wish, given that we are able to exactly measure the velocity  $v_x$ . On the other hand, if we deal with real images, there is always a certain degree of uncertainty in the motion measurement, so that the previous considerations do not hold. This leads us to the issue of considering the sampling theorem in the presence of noisy data (thus generating motion uncertainty). As a consequence of this, we have to know in what way the sampling theorem (as previously described) has to be modified in order to be able to deal with motion uncertainty. Specifically, in the presence of motion uncertainty, it is no longer possible to arbitrarily increase the sampling

interval, without getting aliasing. This establishes a relationship between the maximum sampling interval (in space) (or minimal in the frequency domain) and motion uncertainty.

We know that under the conditions of translational motion (let consider only one-dimensional motion), the Fourier transform of a space-time function has support at the lines passing through the origin, and whose slope is proportional to the velocity of the moving pattern. If we introduce a specific degree of uncertainty for the velocity, then this support will be given by a (one-dimensional) cone, whose aperture is proportional to the uncertainty in the velocity (See Figures 9 and 10).

Considering the case of a band-limited function (with finite support in the frequency domain), we can use polar coordinates to describe its (two-dimensional) variables.

For the angular variable  $\theta$  we have  $\theta = \arctan v_x$  and the radial variable  $r$  is the maximum of  $\sqrt{w^2 + k_x^2}$ . The motion uncertainty  $\Delta v_x$  is given by  $\Delta v_x = (\tan(\theta + \Delta\theta) - \tan \theta)$ , where  $\Delta\theta$  corresponds to the angular aperture of the cone, centered at  $\theta$ . For small values of  $\Delta\theta$ ,  $\delta\theta$ ,  $\Delta v_x$  can be approximated to  $\delta v_x = \sec^2 \theta \delta\theta$ .

If we sample  $f(x + v_x t)$  along the  $x$  direction in intervals of  $\rho_x$  (or  $M_x$  in the frequency domain) (Figure 10), then it is easy to show that, for a fixed motion uncertainty, there exists a minimum value of  $M_x$ ,  $M_x^{\min}$ , such that the adjacent patterns do not overlap.

If we decrease  $M_x$  beyond this threshold, aliasing occurs. This establishes a relationship between  $M_x^{\min}$  and  $\Delta v_x$ , as shown by the following theorem.

**Theorem:** *If we have a band-limited function  $f(x, t)$  describing an uniformly translating pattern, given that its velocity  $v_x$ , which is assumed to be different from zero, is measured within an uncertainty range of  $\Delta v_x$ , then there exists a minimum value for the spatial frequency sampling interval  $M_x^{\min}$  such that no aliasing occurs.  $M_x^{\min}$  is related to  $\Delta v_x$  by*

$$M_x^{\min} = \frac{2r \sin(\Delta\theta/2) \sqrt{1 + \tan^2 \theta}}{\tan \theta + \tan(\Delta\theta/2)},$$

where

$$\Delta\theta = \arctan\left(v_x + \frac{\Delta v_x}{2}\right) - \arctan\left(v_x - \frac{\Delta v_x}{2}\right),$$

$$r = \max \sqrt{w^2 + k_x^2},$$

$$\tan \theta = v_x.$$

**Proof:**

We can observe, from Figure 11 (or Figure 12), that there exists a point  $P$ , in the  $(r, \theta)$  plane, where the adjacent patterns, corresponding to replicas of a (one-dimensional) cone, intersect without overlapping. This point is the solution to the following equations

$$r \cos \theta_1 = M_x^{\min} + d \cos \theta_2 \quad (3.17)$$

and

$$r \sin \theta_1 = d \sin \theta_2, \quad (3.18)$$

where  $\theta_1 = \theta - \Delta\theta/2$  and  $\theta_2 = \theta + \Delta\theta/2$ .

By substituting  $d$ , given by (3.19), into (3.18) we get

$$M_x^{\min} = r \sin \theta_1 (\cot \theta_1 - \cot \theta_2), \quad (3.19)$$

or

$$M_x^{\min} = r \sin(\theta - \Delta\theta/2) \left[ \frac{1}{\tan(\theta - \Delta\theta/2)} - \frac{1}{\tan(\theta + \Delta\theta/2)} \right]. \quad (3.20)$$

Expanding the sine and tangent in (3.21) we get the following expression

$$M_x^{\min} = r(\sin \theta \cos \Delta\theta/2 - \cos \theta \sin \Delta\theta/2) \left[ \frac{1 + \tan \theta \tan \Delta\theta/2}{\tan \theta - \tan \Delta\theta/2} - \frac{1 - \tan \theta \tan \Delta\theta/2}{\tan \theta + \tan \Delta\theta/2} \right], \quad (3.21)$$

which, after some algebra leads to

$$M_x^{\min} = \frac{2r \sin(\Delta\theta/2) \sqrt{1 + \tan^2 \theta}}{\tan \theta + \tan(\Delta\theta/2)}, \quad (3.22)$$

with  $\Delta\theta = \theta_2 - \theta_1$ ,  $\theta_2 = \arctan(v_x + \frac{\Delta v_x}{2})$  and  $\theta_1 = \arctan(v_x - \frac{\Delta v_x}{2})$ .

This concludes the proof.

The theorem shows us that the minimum interval, in the frequency domain, between adjacent sampling points (on the  $k_x$  axis) is bound, nonlinearly, by the degree of motion uncertainty. Consequently, the spatial sampling rate  $\rho_x$  cannot be arbitrarily increased, but depends on the amount of motion uncertainty. Since the spatial sampling rate  $\rho_x$  is the inverse of  $M_x$  ( $\rho_x = \frac{1}{M_x}$ ), and  $M_x$  is bound, by motion uncertainty, to a minimum value  $M_x^{\min}$ ,  $\rho_x$  has a maximum value equal to  $\rho_x^{\max}$ . If  $\rho_x > \rho_x^{\max}$ , we have aliasing of adjacent patterns (cones).

## 4 Conclusion

The extraction of optic flow, via space-time filtering, is given in terms of a collection of filters which are tuned to different orientations in space-time. The space-time Gabor and Cascades of the CS or DOG filters are specially suited for this task because they constitute (space-time) oriented filters. I show that it is possible, in particular, to use the cascaded filter approach of Fleet and Jepson [9] as an energy filter, given that the exponential temporal part of the CS filter is substituted by a (temporal) gaussian.

The space-time filtering approach to the extraction of optical flow is implemented on a sequence of images which are closely displaced in time. The temporal interval between successive images in this sequence corresponds to the (temporal) sampling rate, which as we saw before, is not independent of the spatial sampling rate. In general, we want to use the sequence of images in such a way that we are still able to extract the optical flow, but using the minimum number of images. This means that we have to increase the temporal sampling ratio as much as possible, without getting any aliasing effect. As a consequence of this, we have to ask ourselves what is the upper limit for the temporal

(spatial) sampling rate such that:

1. We still are able to use a filtering approach to extract the optical flow
2. We do not get any aliasing effect.

As shown by the theorem of the previous section, for uniformly translating patterns, the maximum spatial sampling interval is determined by the degree of motion uncertainty. The same can be shown for the case of temporal sampling. If we sample at a lower rate than the minimum amount established by the theorem of section 3, we get aliasing. This answers the second part of the question.

The first part of the question is more difficult to be answered. Just as an illustration, we can mention a problem which bears similarities to the use of a filtering or feature matching approaches to the extraction of optical flow. It is the hypothesis of the existence of two, distinct, processes to detect or extract optic flow in humans [18,19], called short and long range processes. They are studied, in psychophysics, as a phenomenon of apparent motion, which is the capability of the human visual system to be able to interpolate the (spatial) position of moving objects between discrete presentations of successive snapshots of the motion. We can establish a general relationship between short-range and the filtering approach to optical flow, and between long-range and the feature matching approach. The short-range process operates in short temporal intervals (between successive frames - also called inter-stimulus interval ISI, ranging from 50 and 100ms) and angular intervals of 15' or less. The long-range process, on the other hand, can take place even for ISI as long as 400 ms [1], and it works mainly through the matching of features (edges, blobs, etc.), thus operating through the identification of elements in successive frames.

If short and long-range processes in humans are really independent and operate through different mechanisms, it can point out to the possibility that if the filtering and feature matching approaches should bear some resemblance with them, then there should exist a definite borderline between both approaches. In this sense we can say that (space-time) *aliasing* is one criteria by which we can decide upon this problem.



**Acknowledgment**

I would like to thank A. Elfes, T. Kanade, L. Matthies and J. Moura for important discussions.

**References**

- [1] S. Ullman, *The Interpretation of Visual Motion*, 1979, M.I.T. Press, Cambridge, Massachusetts.
- [2] B. K. P. Horn and B. G. Schunk, Determining Optical Flow, 1981, *Artificial Intelligence*, 17, 185.
- [3] E. C. Hildreth, Computations Underlying the Measurement of Visual Motion, 1984, *Artificial Intelligence*, 23, 309.
- [4] B. F. Buxton and H. Buxton, Computation of optic flow from the motion of edge features in image sequences, 1985, *Image and Vision Computing*, 2, No 2, 59.
- [5] D. J. Heeger, Optical Flow Using Spatiotemporal Filters, 1987, to appear in the *International Journal of Computer Vision*.
- [6] E. H. Adelson and J. R. Bergen, Spatiotemporal Energy Models for the Perception of Motion, 1985, *Journal of the Optical Society of America A*, 2 (22), 284.
- [7] A. B. Watson and A. J. Ahumada, A Look at Motion in the Frequency Domain, 1983, *Technical Report 84352, NASA-Ames Research Center*, and Model of Human Visual-Motion Sensing, 1985, *Journal of the Optical Society of America A*, 2 (22), 322.
- [8] R. C. Bolles and H.H. Baker, Epipolar-Plane Image Analysis: A Technique for Analysing Motion Sequences, 1985, *Proceedings: DARPA Image Understanding Workshop*, Miami Beach, Florida, 137.
- [9] D. J. Fleet and A. D. Jepson, On the Hierarchical Construction of Orientation and Velocity Selective Filters, 1985, *University of Toronto Technical Report RBCV-TR-85-8*.
- [10] D. J. Fleet and A. D. Jepson, Spatiotemporal Inseparability in Early Vision: Centre-surround Models and Velocity Selectivity, 1985, *Comput. Intell.* 1, 89.
- [11] R. S. Jasinschi, in preparation.
- [12] W. McC. Siebert, *Circuits, Signals, and Systems*, MIT Press.

- [13] C. E. Shannon, Communication in the Presence of Noise, January 1949, *Proceedings of the I.R.E.*, 10.
- [14] F. H. C. Crick, D. C. Marr and T. Poggio, An Information Processing Approach to Understanding the Visual Cortex, 1980, *Artificial Intelligence Laboratory of M.I.T. Memo No. 557*.
- [15] M. Fahle and T. Poggio, Visual Hyperacuity: Spatiotemporal Interpolation in Human Vision, 1981, *Proc. R. Soc. London. B* 213, 451.
- [16] L. H. Manthies, R. S. Szeliski and T. Kanade, Kalman Filter-based Algorithms for Estimating Depth from Image Sequences, 1987, Computer Science Department (CMU) technical report CMU-CS-87-185.
- [17] N. M. Grzywacs and A. L. Yuille, Massively Parallel Implementations of Theories for Apparent Motion, 1987, *Artificial Intelligence Laboratory and Center for Biological Information Processing - MIT, A.I. Memo No 888 and C.B.I.P. Memo No 016*
- [18] O. J. Braddick, Low-level and High-level Processes in Apparent Motion, 1980, *Phil. Trans. R. Soc. Lond. B* 290, 137.
- [19] S. M. Anstis, The Perception of Apparent Motion Movement, 1980, *Phyl. Trans. R. Soc. London* 290, 153.

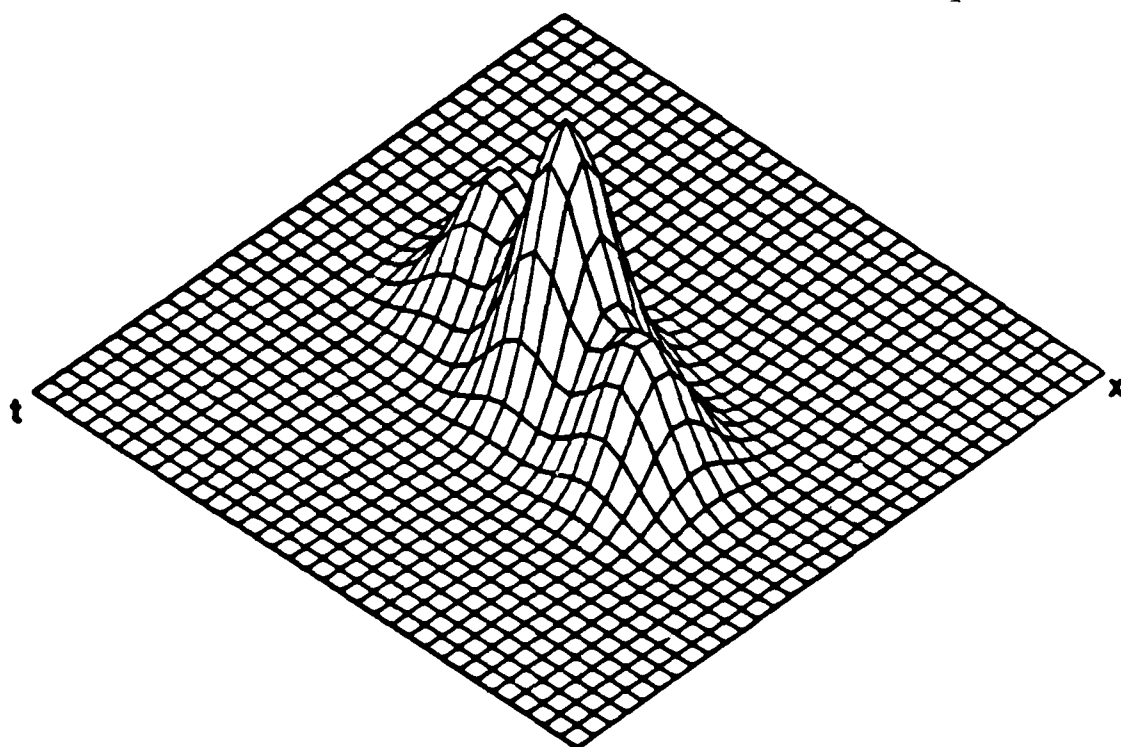


Figure 1 One-layer cascade of space-time DOG filter with  $\xi_x = 0.77$ ,  $\xi_y = 0$ ,  $\tau = 2.89$ .

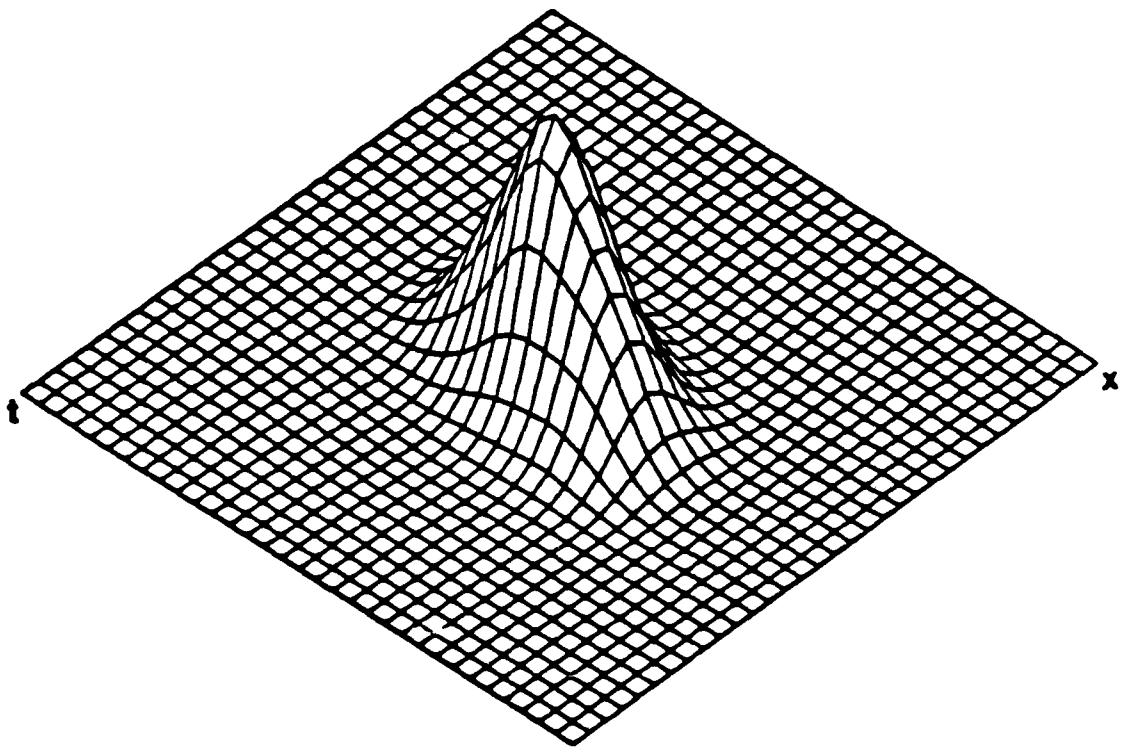
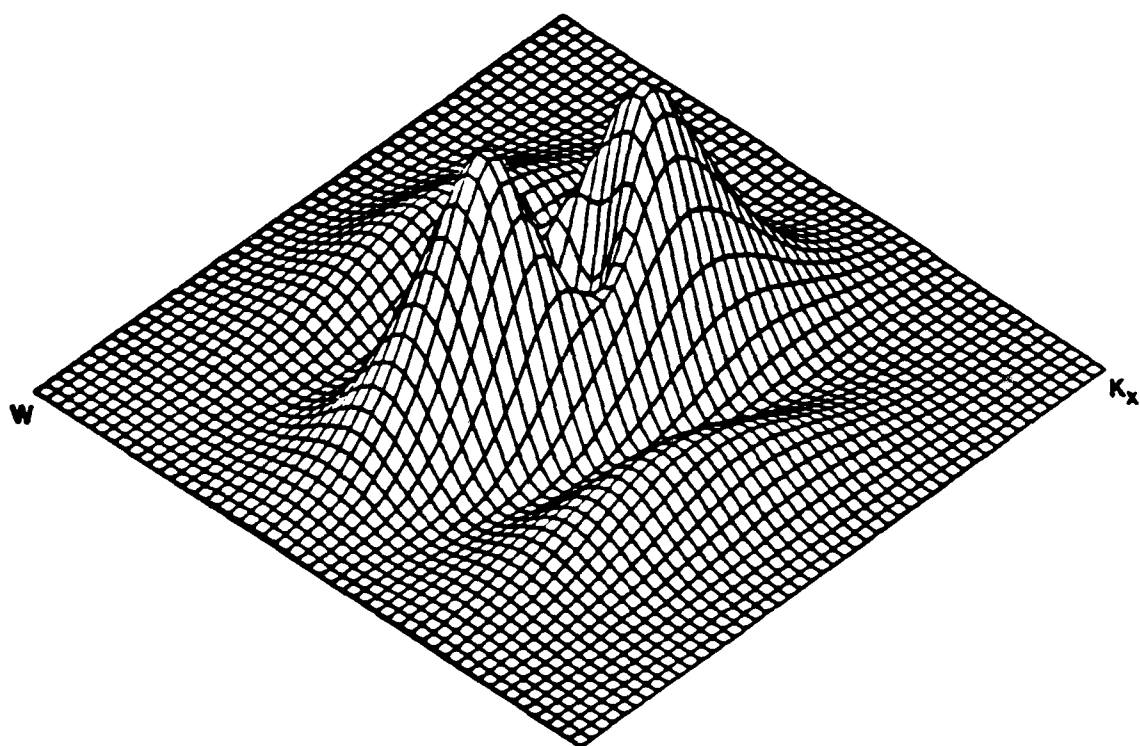
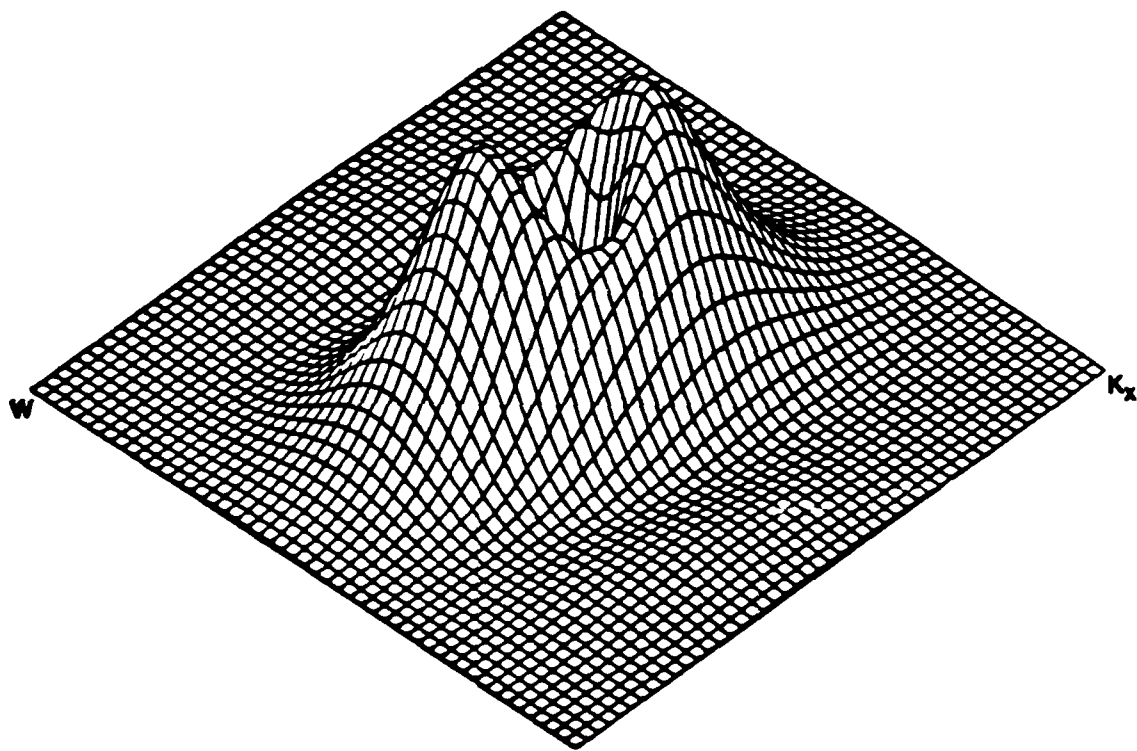


Figure 2 One-layer of a cascade of space-time DOG filter with  $\xi_x = 0.52$ ,  $\xi_y = 0$ ,  $\tau = 1.93$ .



**Figure 3** Fourier transform of the one-layer cascade of space-time DOG filter with  $\xi_x = 0.77$ ,  $\xi_y = 0$ ,  $\tau = 2.89$ .



**Figure 4** Fourier transform of the one-layer cascade of space-time DOG filter with  $\xi_x = 0.52$ ,  $\xi_y = 0$ ,  $\tau = 1.93$ .

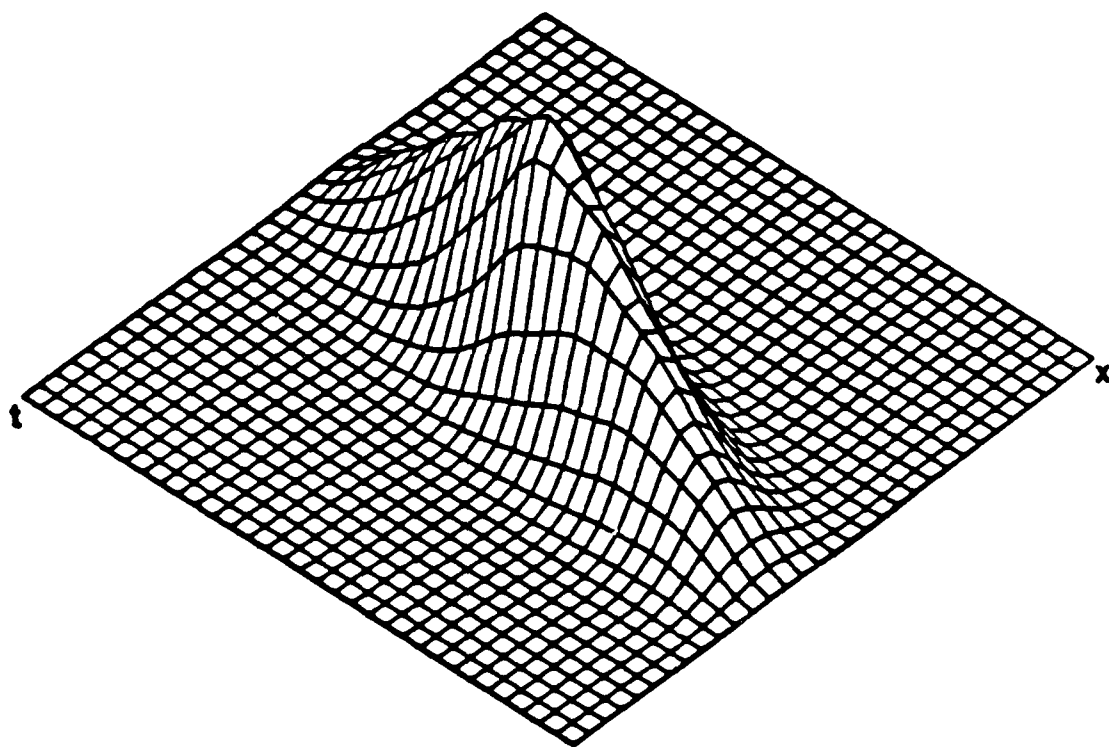


Figure 5 Two-layer cascade of space-time DOG filter with  $\xi_x^1 = 0.52$ ,  $\xi_y^1 = 0$ ,  $\tau^1 = 1.93$ ,  
 $\xi_x^1 = 1.04$ ,  $\xi_y^2 = 0$ ,  $\tau^2 = 3.86$ .



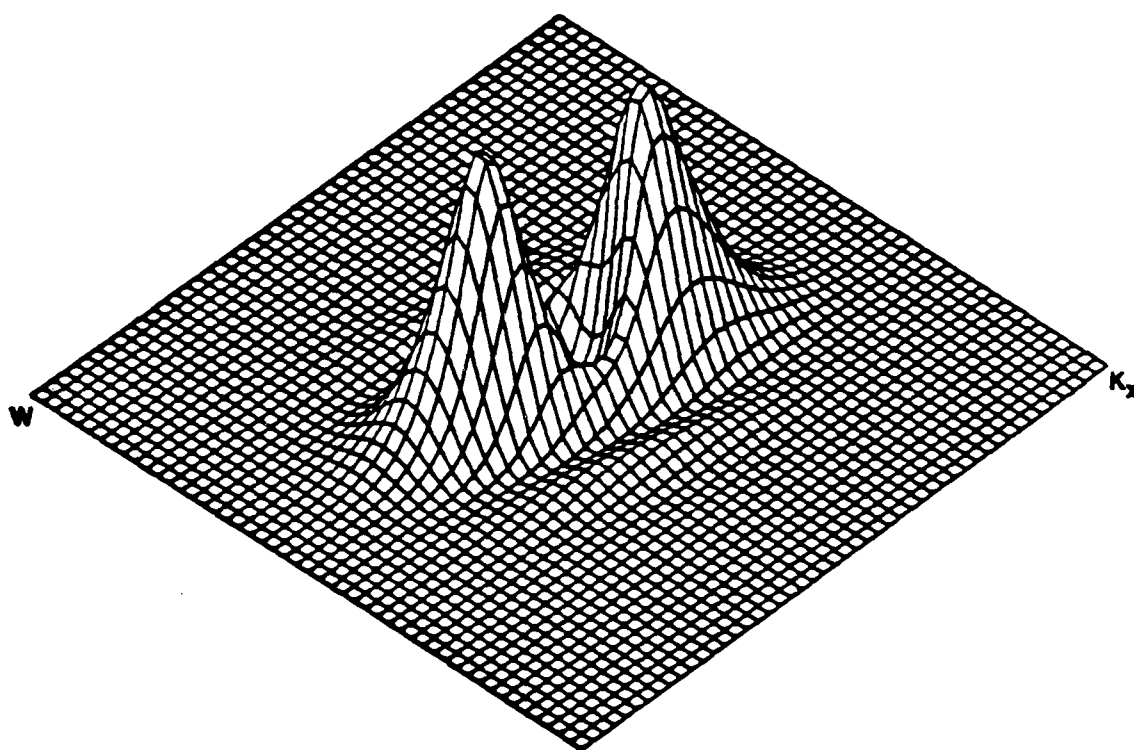


Figure 6 Fourier transform of the one-layer cascade of space-time DOG filter with  $\xi_x^1 = 0.52$ ,  $\xi_y^1 = 0$ ,  $\tau^1 = 1.93$ ,  $\xi_x^2 = 1.04$ ,  $\xi_y^2 = 0$ ,  $\tau^2 = 3.86$ .

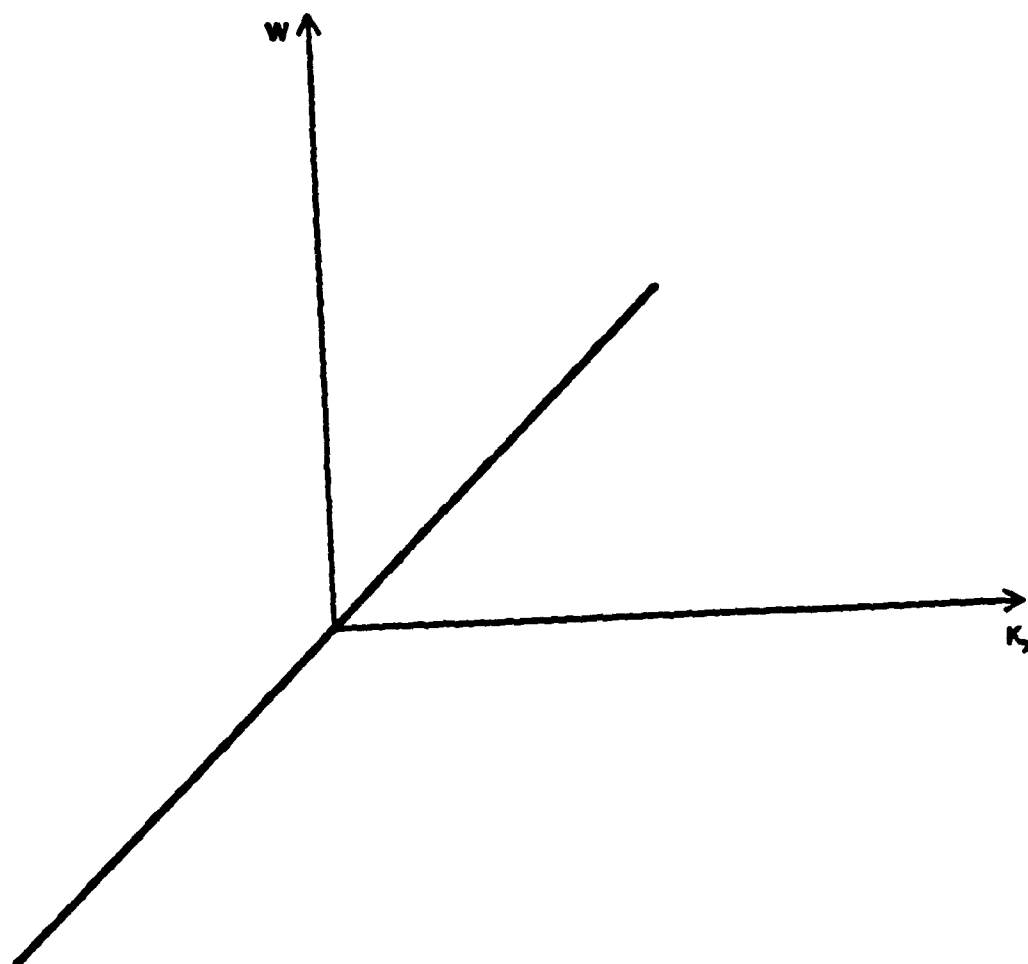


Figure 7 The support, in the frequency domain, for a pattern moving, in one dimension, at an uniform rate. The slope of the segment (band-limited function) of the line is equal to the velocity of the pattern, that is  $w = k_x v_x$ .

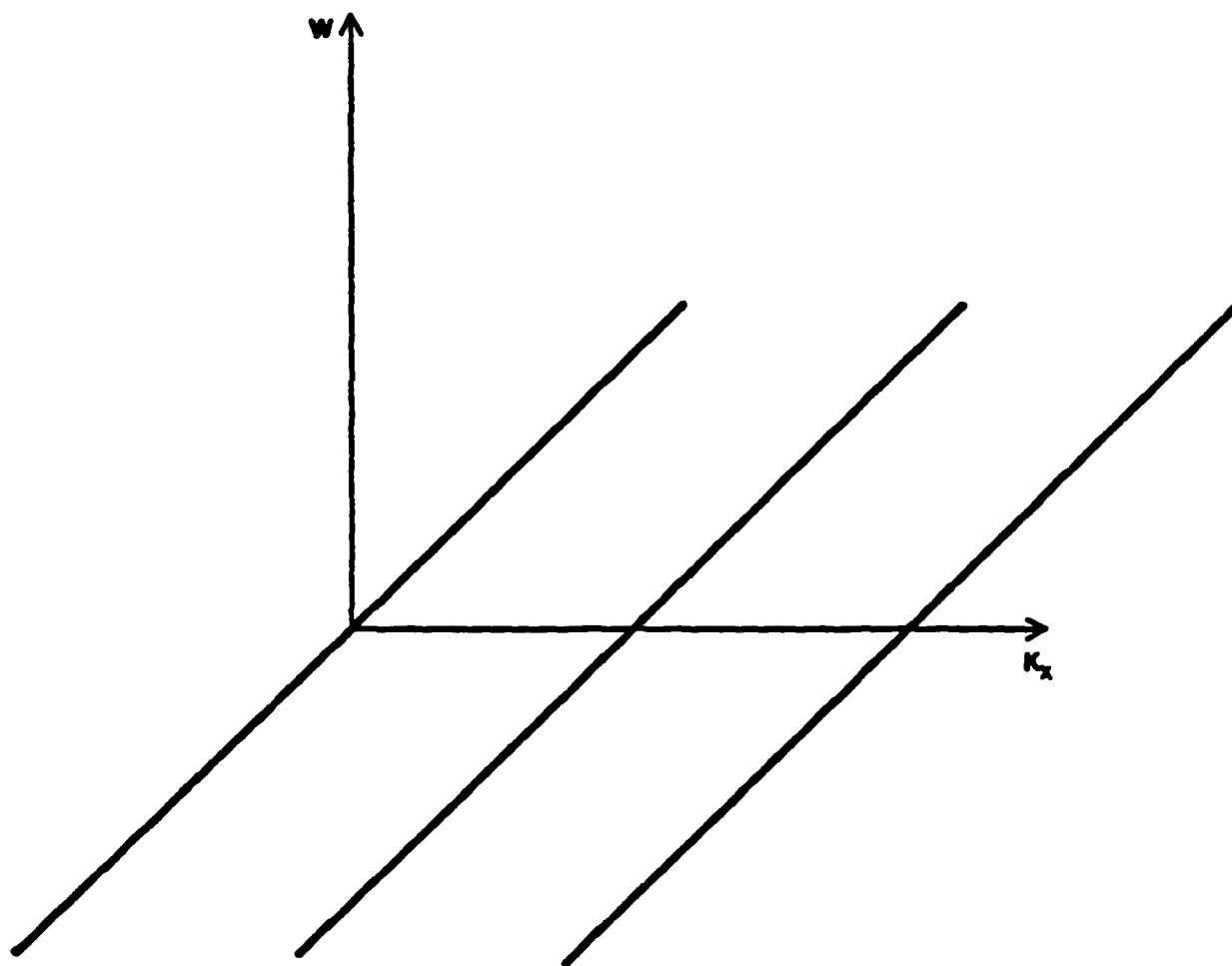
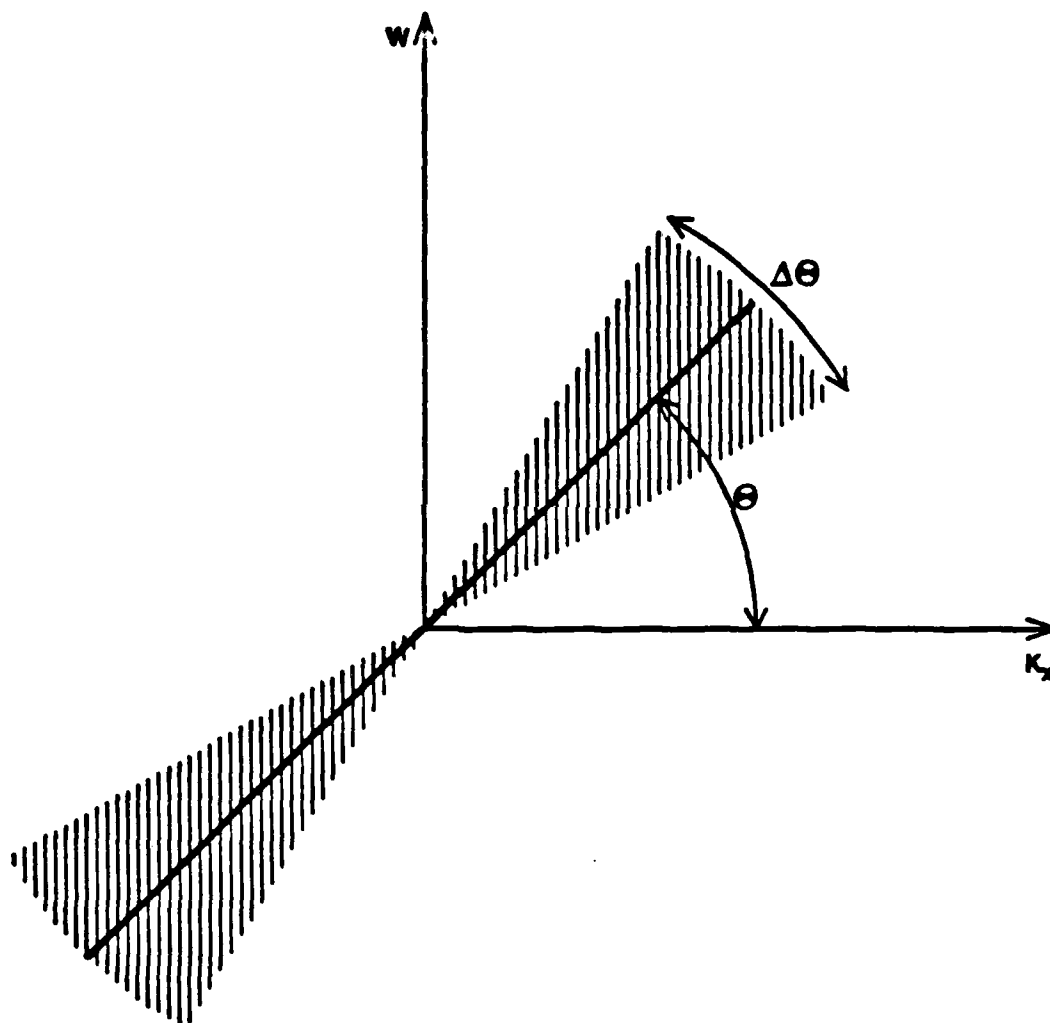
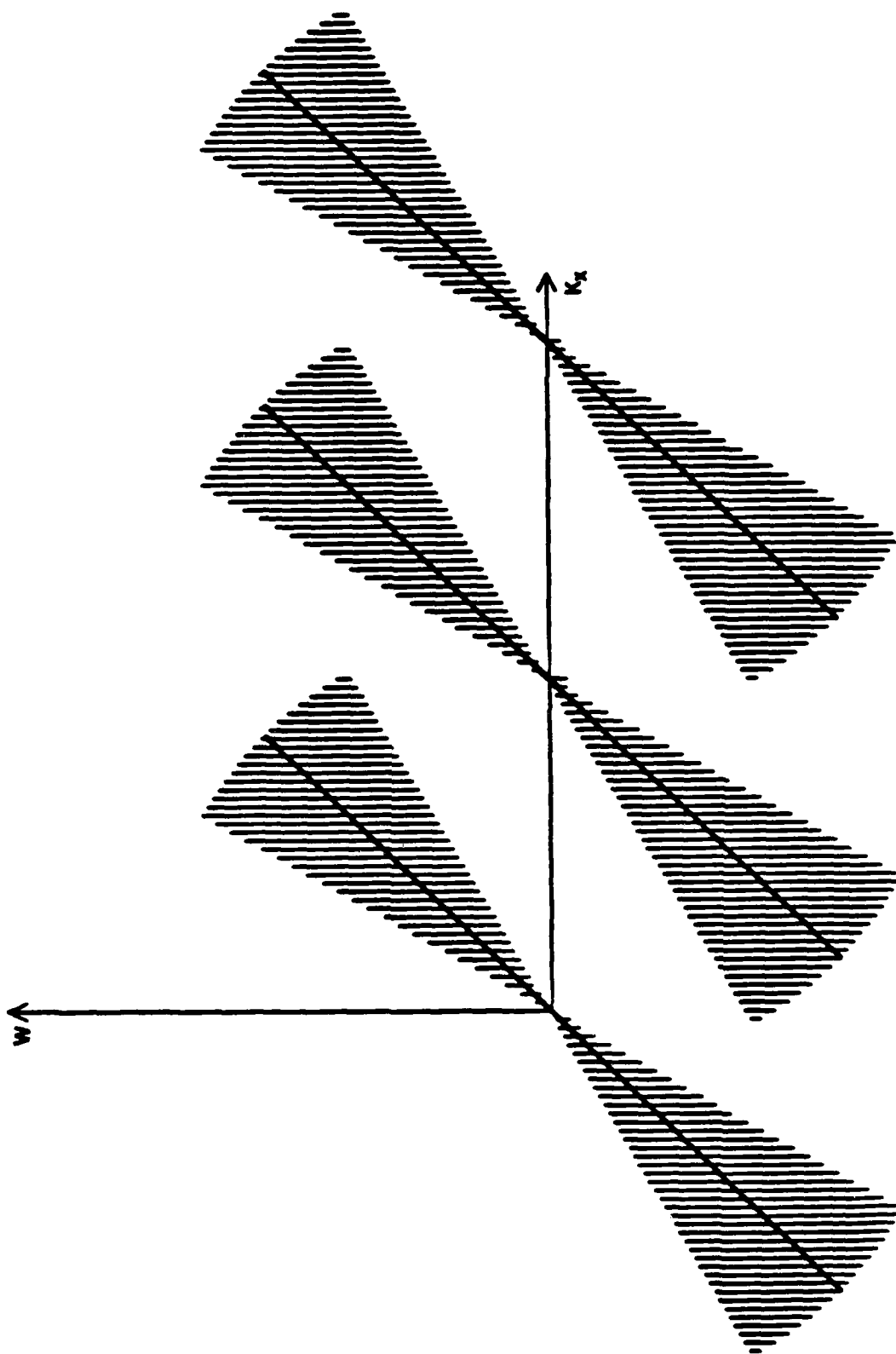


Figure 8 The sampled version of the support of an uniformly translating pattern, as represented by Figure 7. The sampling interval is equal  $M_x$ .



**Figure 9** The support, in the frequency domain, of an uniformly translating pattern whose velocity is measured with a certain uncertainty. The aperture of this cone is equal to this uncertainty.



**Figure 10** The sampled version of the support represented in Figure 9. The sampling interval  $M_x$  is such that the adjacent cones don't overlap.

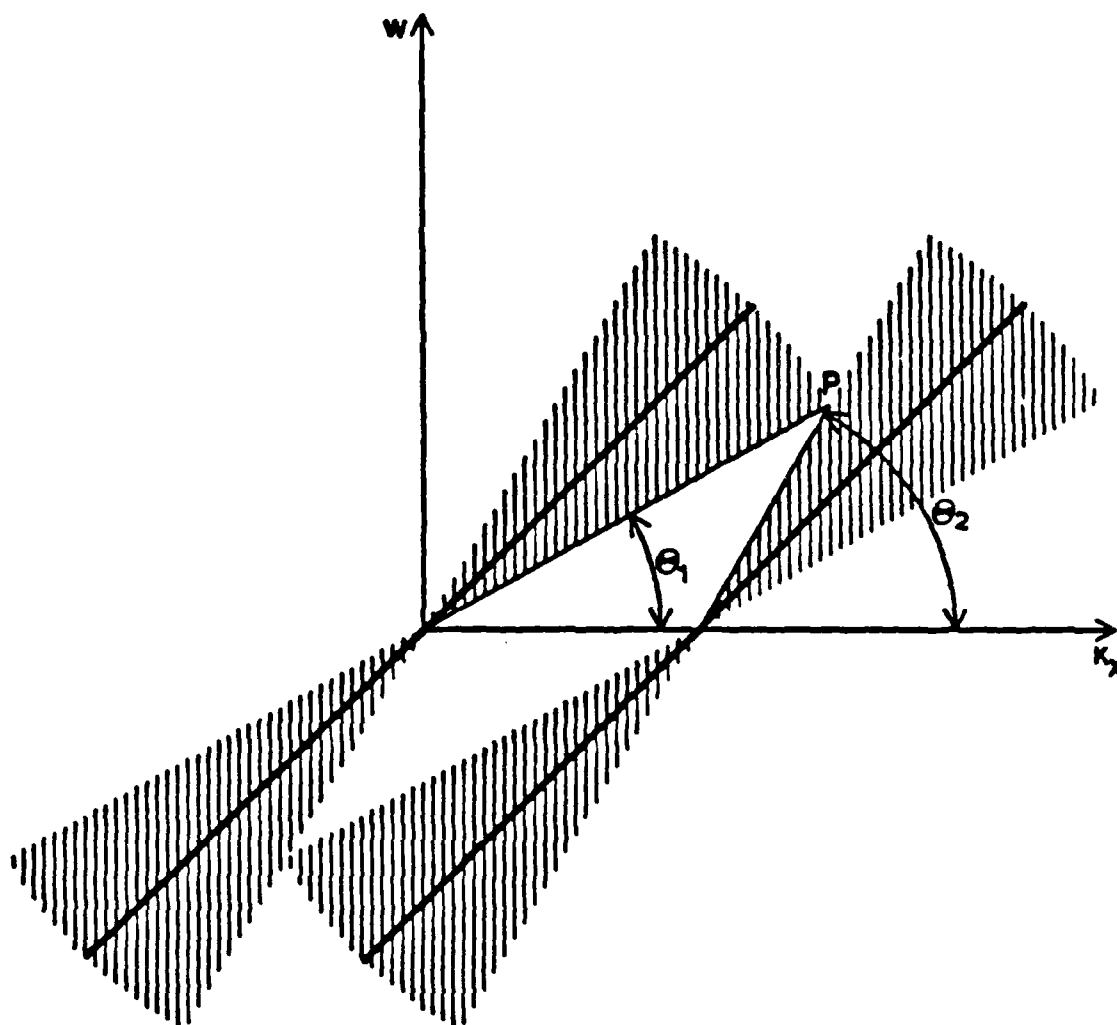


Figure 11 The sampled version of Figure 9 in the case where the sampling rate is such that the adjacent cones touch each other, but without overlapping. The sampling rate  $M_x = M_x^{min}$  is the minimal one such that there doesn't occur aliasing.

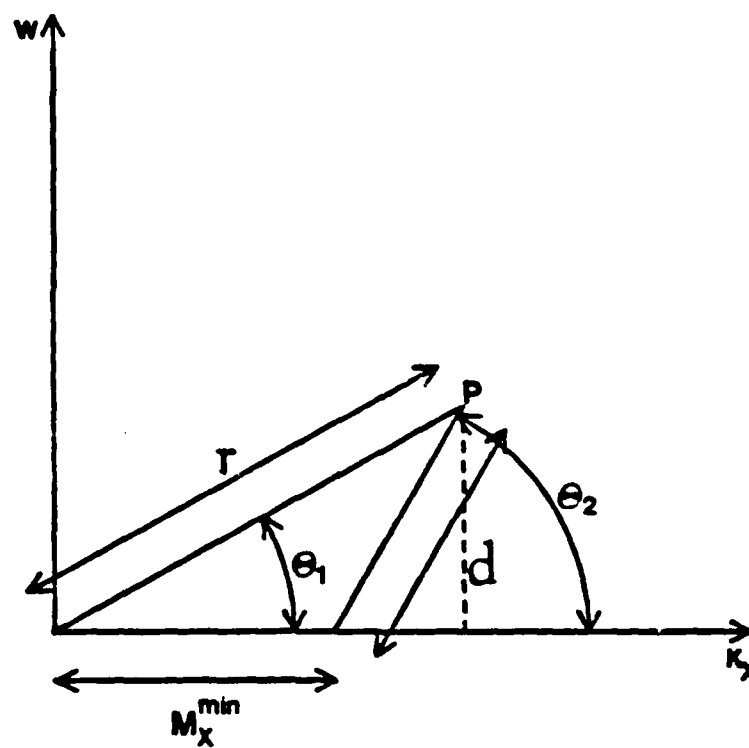


Figure 12 The diagram showing the relevant parameters involved in the proof of the theorem of section 3.